ROBUST RECOGNITION AND CLASSIFICATION OF HERBAL LEAVES

Prabhakar Poudel[1], Shyamdew Kumar[2], Vishal Shaji Philip[3], Pawan Kishore[4], Roopashree S[5]

[1]*Final year B.E. Student, Department of Computer Science, T John Institute of Technology, Karnataka, India*
[2]*Final year B.E. Student, Department of Computer Science, T John Institute of Technology, Karnataka, India*
[3]*Final year B.E. Student, Department of Computer Science, T John Institute of Technology, Karnataka, India*
[4]*Final year B.E. Student, Department of Computer Science, T John Institute of Technology, Karnataka, India*
[5]*Assistant Professor, Department of Computer Science, T John Institute of Technology, Karnataka, India*

Abstract

Our ecosystem has a large variety of plants and they are an essence for other forms of life to flourish. Identifying plants can be simpler if we just identify the leaf of a plant, because each of them has peculiar properties. Plants are often important ingredient for Ayurvedic medicine and also several other modern forms of medication. In this paper we discuss the use of optical method, for the classification of herbs using image processing techniques. Steps needed to execute in an optical method for the herbal leaf recognition include extraction of diverse features, identification of pattern and categorization in sequence. We propose an automation system for the herbal leaf detection to increase consistency, lessen the required time and ease the tagging procedure. A proposed technique also considerably reduces the possibility for human faults, as much of the work is done by the system. This paper hinge on for usage of Support Vector Machine (SVM) for the purpose of classification and Scale Invariant Feature Transform (SIFT) technique to extract features. SIFT features are proved to be invariant to affine transformations, noise and change in illumination. A captured leaf image using a digital camera is pre-processed which gives a binary image as output. The pre-processed image is taken up for extracting essential features, which ultimately is used for the purpose of classification. The projected scheme is suitable to use and effective in terms of cost for the researcher, botanist, and others to discover herbal plants, with high accuracy and effectiveness.

Keywords: *Herbs, Ayurveda, Feature Extraction, Classifier, SIFT, SVM, Image Processing.*

--***--

1. INTRODUCTION

Medicinal history of India goes back a long way, through Ayurveda. At the heart of Ayurvedic medicinal system is plants with medicinal uses or the herbs. Today, a large number of Industries also rely on herbs for manufacturing several different cosmetic and other products. Deforestation and urbanization has given rise to threats these plants are facing and may even come to extinction in near future. This possess a great demand for an optimal mechanism to recognize herbs around for personal or commercial uses. It can also come in handy for the botanists to carry on experiments on them. Often, physical detection can be complicated, consumes time and erroneous. And such knowledge remains inbound to professionals. We propose a computerized mechanism to overcome above limitations. Digital systems have been found to be the best approach to problem solving of any kind, from simple calculator to astronomic computations. Algorithms used here include SIFT feature extraction and SVM classifier. The system is user friendly and no high level knowledge in any field is required to identify herbal leaves.

1.1 Existing System

Historical methods of plant recognition like Cal Floraare, Uconn and Lucid [1, 2, 3] have several shortcomings. Of these mostly include, it takes grayscale inputs and it does not support pre-processing the image. Classification using Probabilistic Neural Network (PNN) was used by Eric You Xu and others [4]. They trained 32 species of leaves for recognition. They achieved an accuracy of up to 90%. Gaussian interpolation and Wavelet Transform (WT) was suggested at [5] for the recognition of leaf using skeleton features. Results exhibited that it is not very much effective than other approaches and could be improved. Experiment taken up by Sandeep in [15] used approaches to classify the medicinal herbs based on extracted features. A survey was taken up on several edge detection methods at [6] and canny edge detection was found better than other approaches. Classification using neural networks for identifying medicinal plans was done at [7] and [8], in which whole plant image was considered and their edges, texture and color was measured. Texture features based classification was done by C. H. Arun [9] to classify medicinal plants. Janani [10] identified the medical plants based on features taken out from leaf using ANN. Automatic approach for plant recognition that employs computer vision and pattern recognition [11] was carried out. Several different approaches have been proposed over the years to classify and recognize plant images. Yet the results have not been much comforting as illumination, rotation and scaling was never effectively dealt with. Affine transformation transfers image into different form and thus preserving co-linearity between points and ratios of distances along a line [12, 13] should be considered like changes in illumination level, rotation, etc.

1.2 Proposed System

We purpose a system that can be used to recognize herbal plants based on the features of their leaves. The feature extraction mechanism used is Scale Invariant Feature Transform and method used for classification is Support Vector Machine. Images are partitioned into training sets and test sets. 70% training images and 30% test image, consisting of both herbal and non-herbal species.

We propose SIFT feature extraction for our system because it is ideally good method for feature extraction where image transformation is an issue. Leaf features are extracted from, and are used for the purpose of storing in case of training images and are compared in case of test images. Larger the number of features extracted, better would be the accuracy in classification. Selection of parameters for each step in extraction determines the number of features used.

SVM classifier takes as input, the features extracted as explained, along with the kernel function and the input parameters that match the kernel function. The stored data is looked for an approximate match between pre-existing features and those given as input. Thus, the image is categorized as either herbal or non-herbal class based on the features extracted. Binary SVM provides with a simple and accurate classifier for such simple tasks.

2. METHODOLOGY

The system is divided into subsystems that perform a predefined functionality at each stage in the program. Subsystems include:

- Acquisition
- Pre-processing
- Extraction using SIFT
- Classification using SVM
- Display

The system architecture can be depicted as shown below:

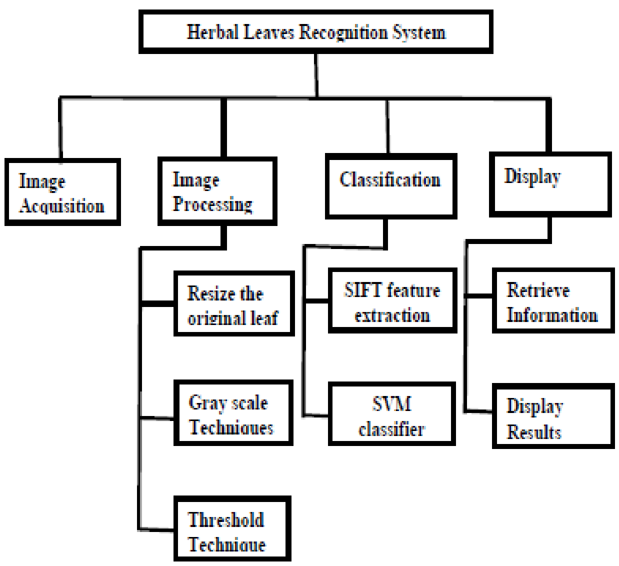

Fig -1: System Architecture

The digital image of a plant leaf, acquired by any of the known means is taken as the input. It is then processed to transform the image into form which is suitable for the process of recognition. The transformed image is the input to the SIFT module, which extracts essential features from the image as a key point descriptor. The descriptor is looked for a match with any of the stored dataset using SVM.

Depending on the outcome of the test appropriate output is displayed.

The proposed system flow diagram is shown below:

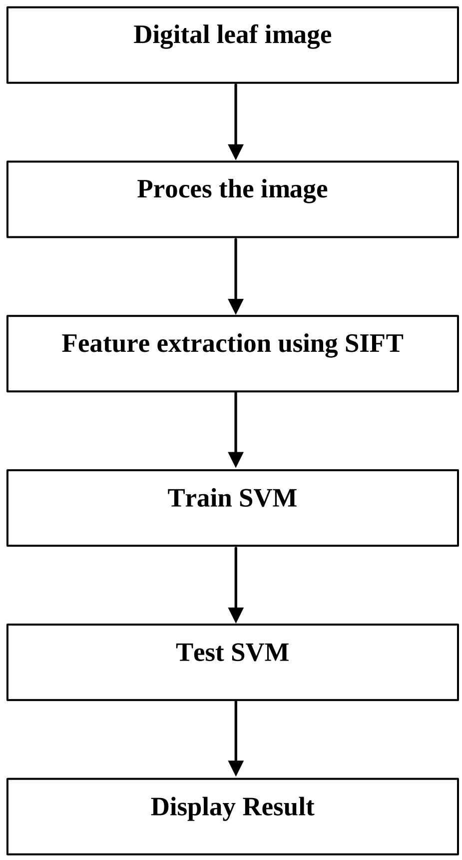

Fig -2: Flow diagram

3. FEATURE EXTRACTION USING SIFT

David Lowe [15] proposed a mathematical algorithmic model for feature extraction, called Scale Invariant Feature Transform, in short SIFT. The steps in SIFT feature extraction involves extraction of key points that is invariant to transformations. The approach here is to filter image in stages that allows different properties to be evaluated. Features extracted are consistent to any alterations in illumination or affine transformation such as scaling and rotation. Co-linearity relation between the points and distance ratio along a line are well preserved in affine transformations.

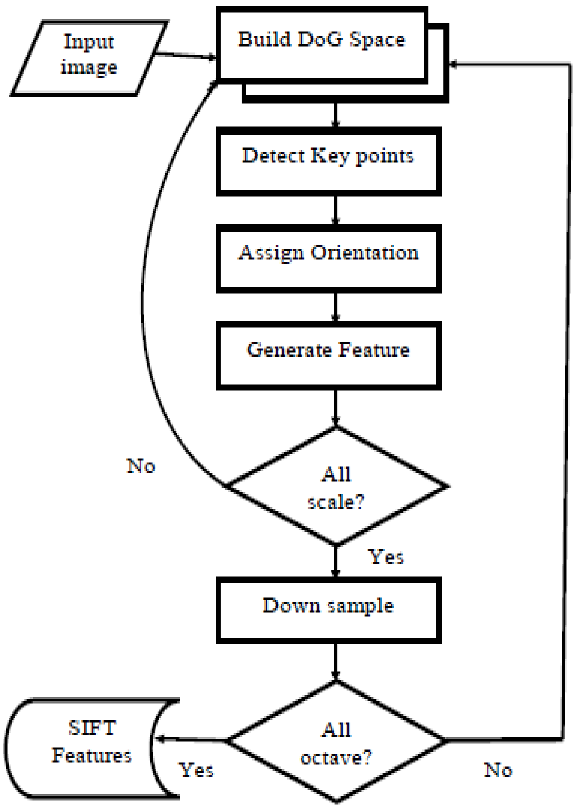

Fig -3: Steps in SIFT process

When compared to approaches such as texture analysis, it is simpler, robust and accurate in the process of extracting features. Here, the features are the key points that can be stored and images can be compared against it. Steps in SIFT can be categorized into the following four stages:

1. Scale-space extrema detection:- Initial steps of SIFT involves searching of all scales in different views of the input image and in different location. Use of continuous function of scale to find stable features is what is called as scale space. Mathematically, it can be executed using Difference-of-Gaussian, consequence is point of interests that remain constant to rotation and scaling transformations.

2. **Key point localization and filtering**:-For all candidate locations there is a detailed model which is fixed that determines scale and location of it. Selection of a key-point is based on the measure of its stability, thus eliminating points on the edge which are not stable.

3. **Orientation assignment**:- Depending on gradient directions of the local image, further orientation is allocated to every key point called beans to make it invariable to rotation location. And all of the future operations to be performed on the data is done using this image which has been transformed relative to location, orientation and scaling for each of the feature and as such provides invariance of transformation.

4. **Key point descriptor**:

The descriptor is thought out at concluding step to attain invariance to the left behind variations. The gradients of local image is measured for a particular region scale about all key point. This is then modified into a representation that permits change in illumination and allows certain levels of local shape distortion.

4. CLASSIFICATION USING SVM

SVM classifier is relatively a new tool for machine-learning that has emerged as very powerful way of training and testing data and resolving binary classification difficulties. The SVM origin goes back to statistical learning assumption where the learning complication was formulated as problem of quadratic optimization. The plane had global optimum and was free of local minimum. The goal here is to locate an optimal separating hyper plane (OSH) or just a Hyperplane that separates two sets of data (+1 and -1).

Finding of OSH by SVM is done by augmenting the margin between those two classes. The idea behind SVM classifier is that first the input data is transformed into higher dimensional space by usage of kernel functions and further construction of OSH between the classes in that transformed space. Those data vectors that lie nearest to the line constructed in that space are called the Support Vectors. SVM approximates a function that classifies data into two diverse classes. Where the partition is linear, into the highdimensional feature space the input vectors are placed by using a non-linear transformation that is based on a regularization factor.

A hyperplane can be considered best for any SVM classifier if it has the peak margin amid the two classes. Here, the border is the maximum width of slab that is parallel to the hyper planes which there are no inner data points.

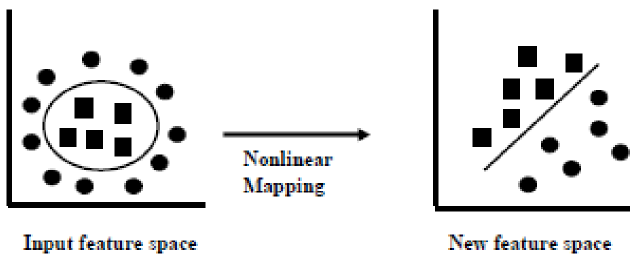

Fig -4: Support Vector Machine classifier

Steps in SVM classification is similar to any supervised learning model. Initial step is to train SVM, next we cross validate the SVM classifier. And then we can use the trained system to classify new data. Also, to obtain better results, we can use any SVM kernel function by tuning the parameter to match the particular kernel function, as in below steps:

- SVM classifier training
- Testing of new data using SVM classifier
- Tuning of an SVM classifier as desired

To construct a nonlinear support vector classifier, the inner product (x, y) is replaced by a kernel function K (x, y)

$$f(x) = \text{sgn}(\sum_{i=1} \alpha_i \gamma_i K(x_i, x) + b)$$

Where, f (x) is a function that determines the membership of x, i = 1, 2… N. The desired class was given as -1 and the others as +1. There are two layers to the functioning of

SVM. First of which selects the basis K(xi, x) during the learning process, from the agreed kernels. And the next layer is responsible for the construction of space linear function.

It is thus comparable of knowing an optimal hyperplane in the given feature space. Support Vector machines can thus be used to build a range of learning machines that employs dissimilar kernel function.

5. CONCLUSION

This paper presents efficient and accurate technique to recognize species of herbal plants based on feature extraction using SIFT and classification using SVM. The technique can detect leaf images with white background. SIFT has the ability of extracting those features that remain unaffected by different transformations. The system is very robust, so is the SVM classifier. Our binary SVM classifier allows us to effectively classify sets of herbal pants from sets of non-herbal ones, features of which is taken from the train dataset. Our approach comprises of four basic steps, starting with pre-processing of the leaf image, extraction of features both for training and testing samples, training the samples and the classification phase. Deploying SVM classifier is as elective in providing the system with a very high recognition rate of more than 90% which can be considered pretty high when compared to other complex or time consuming methods. The performance the system can further be evaluated in terms of the result accuracy and comparison with other methods, and is found to be much flexible.

REFERENCES

[1]. Fu H, Chi Z. IEE Proceedings 2006; 153(6):881-892.
[2]. Xun Liu. Index and Recognition for the Shape Contour of Plant Leaves; Intelligent Systems GCIS WRIGlobal congress; 2009; 436-440.
[3]. Feng Chen, Yun-Feng li. A Multiresolution Segmentation method for Tree Crown Image usingWavelets; International Conference on Wavelet Analysis and Pattern Recognition; 2007; 1556-1559.
[4]. Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, I-Fan Chang, Qiao-Liang Xiang. ALeaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network; IEEEInternational Symposium on Signal Processing and Information Technology; 2007.
[5]. XiaoGu, Ji-Xiang Du, Xiao-Feng Wang. Leaf Recognition Basedon the Combination of WaveletTransform and Gaussian Interpolation; International Conference on Intelligent Computing; 2005;3644: 253-262.
[6]. "Canny Edge Detection" 09gr820, March 23, 2009.
[7]. Manisha Kaushal, Arjan Singh, Baljit Singh, "Adaptive thresholding forEdge Detection in Gray Scale Images", international Journal ofEngineering Science and Technology", Vol. 2(6), pp. 2077-2082, Jan.2010.
[8]. Raman Mani, Dr. Himanshu Aggarwal, "Study and Comparison ofVarious Image Edge Detection Techniques", international Journal of image Processing (lJIP), Vol.3 (I), pp. 1- 12, Jan. 20 II.
[9]. C.H Arun, W.R Sam Emmanuel, D. Christopher Dhurairaj, "TextureFeature Extraction for Identification of Medicinal Plants and Comparison of Different Classifiers", international Journal of Computer Applications (0975 - 8887), Volume 62(12), Jan. 2013.
[10]. Janani R, Gopal A, "Identification of selected medicinal plant leaves using Image features and ANN", IEEE international Conference on Advanced Electronic Systems (ISAES), pp. 238- 242, Sept. 2013.
[11]. Pan J, He Y. Recognition of plants by leaves digital image and neural network; IEEE InternationalConference on Computer Science and Software Engineering; 2008
[12]. De Villiers M. The affine invariance and line symmetries of the conics; Australian Senior MathematicsJournal; 1993
[13]. Yaglom Y. Geometric Transformations III; Mathematical Association of America (MAA), Washington;1973; 9-20.
[14]. Lowe David G. Object Recognition from Local Scale-Invariant Features; Proceedings of the SeventhIEEE International Conference on Computer Vision; 1999;
[15]. E. Sandeep Kumar. Recognition of Medicinal Plants based on its Leaf Features; Volume 327 of the series Lecture Notes in Electrical Engineering pp 99-113; 2015

BIOGRAPHIES

Prabhakar Poudel, Eighth Semester, Department of Computer Science and Engineering, T John Institute of Technology, Bengaluru

Shyamdew Kumar, Eighth Semester, Department of Computer Science and Engineering, T John Institute of Technology, Bengaluru

Vishal Shaji Philip, Eighth Semester Department of Computer Science and Engineering, T John Institute of Technology, Bengaluru

Pawan Kishore, Eighth Semester, Department of Computer Science and Engineering, T John Institute of Technology, Bengaluru

Prof. Roopashree S received her BE degree in 2001 from Department of Computer Science and Engineering, VTU and MTech in Computer Science and Engineering, VTU. She is currently persuing her PhD in Department of Computer Science and Engineering, VTU in Machine Learning and Artificial Intelligence. She has published various papers in National and International journals.